第65卷 第2期 天 文 学 报 Vol. 65 No. 2
2024年3月 ACTA ASTRONOMICA SINICA Mar., 2024

doi: 10.15940/j.cnki.0001-5245.2024.02.005

EfficientNetV2-S-Triplet7: 一种改进的星系形态学分类算法*

仲峥迪‡ 屠良平† 冯雪琦 李 娟 李 馨

(辽宁科技大学理学院 鞍山 114051)

摘要 星系的形态与星系的形成和演化息息相关, 其形态学分类是星系天文学后续研究的重要一环. 当前海量天文观测数据的出现使得天文数据自动分析方法越来越得到重视, 针对此问题, 利用先进的深度学习骨干网络EfficientNetV2, 分析不同的注意力机制类型和使用节点对网络性能的影响, 构建了一种命名为EfficientNetV2-S-Triplet7 (即在EfficientNetV2-S stage7的1×1卷积层后加入Triplet模块)的改进算法模型来实现星系形态学的自动分类. 使用第二期星系动物园(Galaxy Zoo 2, GZ2)中超过24万张的测光图像作为初始数据进行实验测试. 在对数据进行预处理时采取了尺寸抖动、翻转、色彩畸变等图像增强手段来解决图像数量的不平衡问题. 在同一系列经典和前沿的深度学习算法模型AlexNet、ResNet-34、MobileNetV2、RegNet进行对比实验后, 得出EfficientNetV2-S-Triplet7算法在分类准确率、查全率和F1分数等指标上具有最好的测试结果. 在9375张测试图像中的3项指标值分别可达到89.03%、90.21%、89.93%, 查准率达到89.69%, 在其他模型中排在第3位. 该结果表明将EfficientNetV2-S-Triplet7算法应用于大规模星系数据的形态学分类任务中有很好的效果.

关键词 技术: 图像处理, 方法: 数据分析, 方法: 分类, 星系: 结构

中图分类号: P157; **文献标识码**: A

1 引言

宇宙中的星系呈现出各种形状、颜色和大小. 这些特性表明了它们的年龄、形成条件以及在许多环流过程中与其他星系的相互作用[1]. 为了理解星系的形态与其形成机制之间的关系, 天文学家展开了对星系的形态学分类的研究[2]. 应用最广泛的有关星系的分类是来自Hubble于1926年提出的分类系统, 即著名的星系分类"音叉图"[3]. 早期哈勃的分类系统是基于人眼进行观察分类的, 而且分类的程度也比较粗糙. 基于"Hubble音叉图", 后来的天文学者们对星系形态学分类进行了发展, 提出了多种更为细化的目视分类系统, 如德沃古勒系统[4]和叶凯士系统[5].

由于基于目视的星系形态分类方法易受人类判断的主观影响, 且无法完成对海量星系数据的分类, 所以学者们发展了新型的基于计算机的分类系统, 包括模型化分类系统和非模型化分类系统. 基于星系形态不同, 星系的面亮度分布各有不同的特点, 用不同的经验律对星系的面亮度轮廓进行拟合的分类系统称为模型化分类系统. 但模型化分类系

2022-12-15收到原稿, 2023-02-23收到修改稿

*国家自然科学基金项目(U1731128)资助

†tuliangping@ustl.edu.cn

‡1532800873@qq.com

统也有其自身的局限性: 它依赖于假定的函数形式, 且没有考虑到星系可能存在的其他内部结构等. 为了克服这些问题, 又提出了非模型化分类系统, 这是使用聚集度指数[6]、基尼指数[7]等结构参量对星系形态进行分类的方法. 相对于传统的方法, 非模型化分类系统的优点是: 不需要提前假定星系光度分布函数, 可以完全脱离人工划分, 使用计算机自动分类, 也能适用于大样本数据分类研究. 此外, 表征星系形态的结构参量具有丰富的物理内涵, 星系的形成历史和演化与其密不可分, 因此它们在不同红移的形态学分类研究中的应用越来越广泛[2].

随着科学技术的飞速发展, 观测技术也取得了较大进步, 逐步建立了以斯隆数字巡天(Sloan Digital Sky Survey, SDSS)[1]为代表的大天区星系图像巡天计划, 观测得到的星系数据增长十分迅猛. 因此, 在星系形成和演化的物理研究方面, 使用人工对海量星系数据进行分类显得不再可行. 随着机器学习迅速成为自动化数据密集型处理过程的首选工具, 许多天文工作者也考虑将其应用到天文数据处理中. 例如, 人工神经网络(Artificial Neural Network, ANN)[8]、朴素贝叶斯(Naive Bayes)[9]、决策树(Decision Tree)[10]和局部加权回归(Locally Weighted Regression)[11]已经在早期工作中应用于相对较小数据集的星系分类任务中. 后来, 大规模数据的发展为机器学习在星系分类任务中的应用提供了驱动力. Cheng等[12]先将10种传统机器学习方法应用于星系动物园(Galaxy Zoo, GZ)数据集中, 再使用监督学习校正分类错误星系的标签, 在椭圆星系和旋涡星系的二分类任务中达到了超过0.99的准确率. Zhang等[13]引入了一种基于少样本学习(Few-shot learning)策略的SC-Net (Self-Calibrated Convolutions), 该模型与AlexNet相比, 在仅含1000张星系图像的训练集上的分类准确率提高了21%.

深度学习作为更复杂的机器学习模型, 在大规模数据中表现更佳. 近年来, 越来越多的天文学者和计算机视觉工作者参与到基于深度学习算法的星系形态学分类方法的研究中. Dieleman等[1]首次提出进行训练使用的模型是卷积神经网络, 使用的训练集也超过了5万张, 最终取得的均方根值为0.07492. 同时, Dieleman[1]还发现了星系图片的旋转和翻转不变性. Kim等[14]利用SDSS DR12 (Data Release 12)和CFHTLenS (Canada-France-Hawaii Telescope Lensing Survey)[2]中的星系、恒星图像作为数据集, 提出了一个可以用于二分类的深度为11层的卷积神经网络, 测试集上恒星和星系的分类准确率分别达到了99.52%和99.48%. 戴佳明等[15]基于旋涡星系、侧向星系等5类星系形态分类问题, 设计了ResNet-26算法, 这是一个改进的深度残差网络的分类框架, 在测试集上准确率达到了95.2%. Goyal等[16]采用基于卷积神经网络的顺序分层结构和数据增强技术, 将星系分为椭圆星系、透镜星系和旋涡星系3大类. 模型的平均训练准确率达到了90.2%, 平均验证准确率为88.3%. 王林倩等[17]利用SDSS DR16、星系动物园2 (Galaxy Zoo 2, GZ2)[3]、EFIGI (Extraction de Formes Idealisées de Galaxies en Imagerie)[4]目录中星系的测光图像, 提出了一种名为GMC (Galaxy Morphological Classification)的星系形态学分类方法, 在旋涡星系、椭圆星系、透镜星系以及不规则星系分类问题上达到98.93%的准确率. 艾霖嫔等[18]利用戴佳明等[15]使用的数据集, 基于EfficientNet-B5模型, 达到了96.76%的分类准确率.

然而, 近年来学者的分类问题研究多致力于Kaggle平台Galaxy Zoo-The Galaxy Challenge[5]的数据集, 其星系数量约有14万. 本文则针对GZ2中超过24万的数据量, 基于EfficientNetV2-S[19]网络对其进行改进, 嵌入Triplet[20]注意力模块, 提高了模型特征提取的能力. 并结合尺寸抖动、翻转、色彩畸变等数据增强技术, 提高数据集的多样性, 降低了过拟合, 使得在椭圆星系、有核球侧向星系、无核球侧向星系、旋涡星系、棒旋星系和扰动星

[1]https://www.sdss.org
[2]https://www.cfhtlens.org/
[3]https://www.galaxyzoo.org/
[4]https://www.efigi.org
[5]https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge

系6分类任务上的准确率均有所提高.

2 数据

本文采用的实验数据全部来自GZ2. 本节主要介绍本文使用的数据集及其获取、数据预处理过程.

2.1 GZ2数据集

GZ2是一个要求参与者描述基于彩色图像的星系形态的在线众包项目, 其中超过1600万的形态学分类来自SDSS释放的约30万个星系数据. 参与者会被问到各种各样的有关星系形态学的问题, 这些问题作为标准形成了一个决策树. 分类方案中共设计了11个问题(T01–T11)和37个答案(如T01.1表示第1个问题的第1个答案), 这37个答案分别对应37个独立的类别. 事实上, 其中一些类别有许多共同属性. 由于决策树的结构, 每个参与者只回答了每个分类问题的一个子集. 当许多参与者对同一图像进行分类时, 他们的答案被聚合成一组关于整个决策树的加权投票分数. 这些投票分数用于估计每个答案的置信水平, 并指出用户在对图像进行分类时所经历的困难程度. 共有超过50万人参与了该分类工作, 每张图片都由40–50人独立分类[21].

GZ2中的数据是尺寸为$424 \times 424 \times 3$像素的RGB图片, 其测光数据来自于SDSS. 事实上, SDSS得到的是含有u、g、r、i和z这5个波段的原始数据, 但是u和z波段大多是近红外和近紫外, 包含的有用信息比较少, 因此在合成相对应的RGB图像时, 一般只选用g、r、i这3个波段的信息.

来自GZ2项目的数据已经被广泛用于关于星系结构、形成和演化的研究. 将GZ2的形态与来自权威天文学专家和自动分类的较小样本进行比较, 结果显示出了高度的一致性, 进而证明了众包注释的准确性. 所以本文采用的GZ2数据具有较为准确的分类标签.

2.2 数据准备

表1中为筛选6类星系图像数据的阈值标准, 6类星系的类别标签0、1、2、3、4、5分别代表椭圆星系、有核球侧向星系、无核球侧向星系、旋涡星系、棒旋星系和扰动星系. (d)表示扰动星系. f_{smooth}表示一张图片是平滑星系的概率; $f_{\mathrm{in\ between}}$表示一张图片是中间星系的概率; $f_{\mathrm{features/disk}}$表示一张图片有盘状结构的概率; $f_{\mathrm{edge\text{-}on,\ yes}}$表示一张图片是侧向星系的概率; $f_{\mathrm{bulge,\ yes}}$表示一张图片有核球结构的概率; $f_{\mathrm{bulge,\ no}}$表示一张图片无核球结构的概率; $f_{\mathrm{edge\text{-}on,\ no}}$表示一张图片不是侧向星系的概率; $f_{\mathrm{a\ bar\ through\ the\ centre,\ no}}$表示一张图片中间没有棒状结构的概率; $f_{\mathrm{spiral,\ yes}}$表示一张图片是旋涡星系的概率; $f_{\mathrm{a\ bar\ through\ the\ centre,\ yes}}$表示一张图片中间有棒状结构的概率. N1代表依据阈值标准直接在GZ2数据中选取出的星系图像数量, N2代表删除了带有重复标签的星系图像数量, N3代表进行数据增强后星系图像数量.

表1的阈值标准来自GZ2的阈值选取规则, 官方未给出的阈值用0.8, 无核球侧向星系适当放宽至0.5, 同时还要满足所有条件概率值相乘大于0.5. 比如, 对于有核球侧向星系, 筛选的是所有满足$f_{\mathrm{features/disk}} > 0.43$、$f_{\mathrm{edge\text{-}on,\ yes}} > 0.602$、$f_{\mathrm{bulge,\ yes}} > 0.8$和$f_{\mathrm{features/disk}} \times f_{\mathrm{edge\text{-}on,\ yes}} \times f_{\mathrm{bulge,\ yes}} > 0.5$ 的交集图片.

筛选出的77045张图片中, 有329张星系图片既分在了旋涡星系又分在了扰动星系, 有259张既分在了棒旋星系又分在了扰动星系, 这不符合深度学习数据集的要求, 因此删掉这1176张图片, 共得到75869张星系图片. 选取后椭圆星系、有核球侧向星系、无核球侧向星系、旋涡星系、棒旋星系和扰动星系图片数量分别为22897、9157、2419、23992、15888、1516张. 我们发现, 无核球侧向星系和扰动星系的数量较少, 样本类别数量不平衡, 不利于模型充分学习各个类别的特征信息. 所以采用数据增强的方式扩充无核球侧向星系和扰动星系的数量. 这两类样本(2和5)数量分别由2419和1516扩充到9676和12128个, 6种类别的总样本数由75869扩充到93738个. 本文采用的数据增强方式如下:

旋转: 旋转不变性是星系图像特有的一种性质, 利用图像的该性质对图像在$-30°$到$30°$范围内进行随机旋转.

翻转: 水平旋转和垂直翻转. 这样就消除了模型对于不同方向的星系分类可能存在的差异.

色彩畸变: 包括改变亮度、对比度、饱和度和色调, 以便严格保留星系的形态信息.

将93738张干净样本按8:1:1的比例在各类别中随机分层抽取, 划分为训练集、验证集和测试集. 表2详细展示了训练集、验证集和测试集中每一类星系图像的数目, 为了保证实验的科学性和准确性, 3个数据集中各类别星系图像数目是同比例分布的.

表 1 星系样本数据选取标准

Table 1 Selection criteria for sample data of galaxies

Class	Task	Threshold setting	N1	N2	N3
0	T01.1	$f_{\rm smooth} > 0.469$	22897	2289	22897
	T07.2	$f_{\rm in\ between} > 0.8$			
1	T01.2	$f_{\rm features/disk} > 0.430$	9157	9157	9157
	T02.1	$f_{\rm edge\text{-}on,\ yes} > 0.602$			
	T09.1&.2	$f_{\rm bulge,\ yes} > 0.8$			
2	T01.2	$f_{\rm features/disk} > 0.430$	2419	2419	9676
	T02.1	$f_{\rm edge\text{-}on,\ yes} > 0.602$			
	T09.3	$f_{\rm bulge,\ no} > 0.5$			
3	T01.2	$f_{\rm features/disk} > 0.430$	24321	23992	23992
	T02.2	$f_{\rm edge\text{-}on,\ no} > 0.715$			
	T03.2	$f_{\rm a\ bar\ through\ the\ centre,\ no} > 0.8$			
	T04.1	$f_{\rm spiral,\ yes} > 0.619$			
4	T01.2	$f_{\rm features/disk} > 0.430$	16147	15888	15888
	T02.2	$f_{\rm edge\text{-}on,\ no} > 0.715$			
	T03.1	$f_{\rm a\ bar\ through\ the\ centre,\ yes} > 0.8$			
	T04.1	$f_{\rm spiral,\ yes} > 0.619$			
5	(d)		2104	1516	12128
Total			77045	75869	93738

表 2 划分后6种类别的星系样本统计

Table 2 Statistics of six categories of the galaxy sample after partition

Data type	0	1	2	3	4	5	Total
Training set	18317	7325	7740	19193	12710	9702	74987
Validation set	2290	916	968	2399	1589	1213	9735
Test set	2290	916	968	2400	1589	1213	9736
Data set	22897	9157	9676	23992	15888	12128	93738

2.3 数据预处理

天文图像含有很多冗余的信息, 选取后的训练集、验证集、测试集星系图片不能直接进行训练, 输入模型前需要进行数据预处理. 为了保证验证和测试样本的真实性, 验证集和测试集的数据预处理方式较为简单, 与训练集的处理方式有所不同.

训练集图像预处理流程如图1所示, 从左到右依次为原图、裁剪、缩放、随机翻转生成的图片, 具体说明如下.

(1)裁剪: 星系图像数据带有大量天空背景, 我们感兴趣的目标位于中间, 裁剪可以消除周围噪声. 本文采用尺寸抖动的方法, 将424 × 424 pixel尺寸抖动到[220, 260], 即将图片像素裁剪到220到260的任意一个整数. 这种策略可以使每张训练图像经过每次迭代, 都能获得随机的裁剪尺寸, 从而大大增加训练集的数量, 有效避免过拟合.

(2)缩放: 降低图像分辨率. 将图片缩放到90 × 90 pixel大小, 这样虽然减少了图像中包含的细节, 但是可以保证不丢失数据的任何特征.

(3)随机翻转: 以0.5的概率对图像进行随机水平翻转和随机垂直翻转. 这样就消除了模型对不同方向的星系解释上可能存在的差异.

验证集和测试集处理步骤相对简单. 先中心裁剪到240 × 240 pixel, 再缩放到90 × 90 pixel. 将上述训练集、验证集和测试集的RGB图像先转为张量, 再分别按通道进行标准化, 即先减去平均值再除以标准差, 其中通道分别为H (Height)、W (Width)、C (Channel).

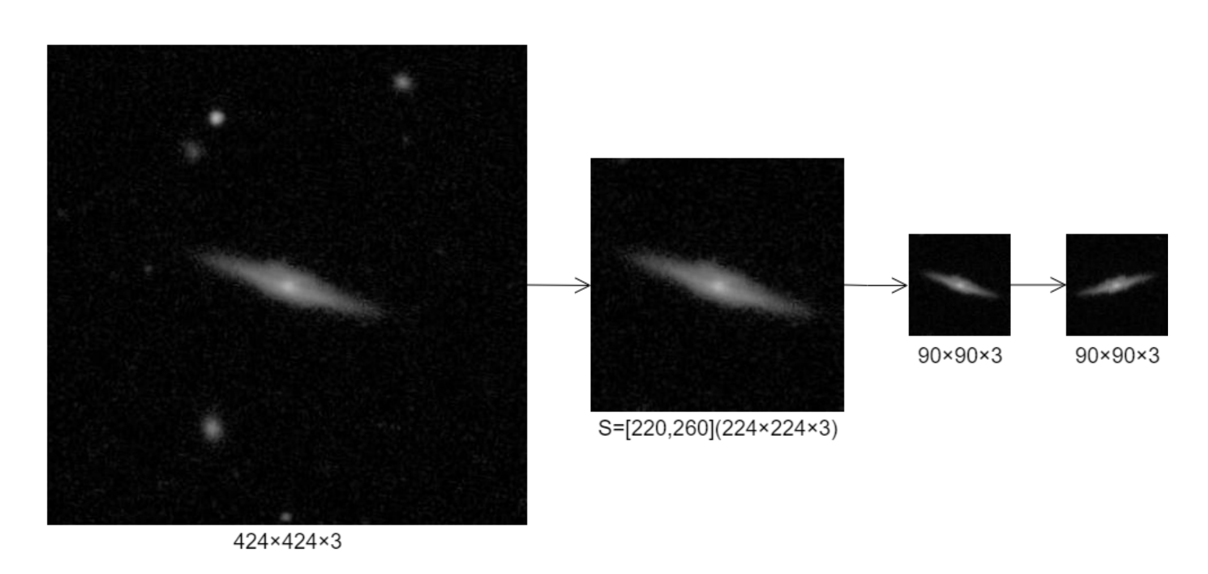


图 1 训练集中一个有核球侧向星系的预处理过程(asset_id = 10508)

Fig. 1 Preprocessing of an edge-on with bulge galaxies in the training set (asset_id = 10508)

3 模型架构

卷积神经网络(Convolutional Neural Networks, CNN)通常包含卷积层、池化层和全连接层, 常被设计用于处理图像数据[22], 是深度学习的代表算法之一[23]. 本节介绍实验中使用的一些主要模型架构, 3.1节介绍了高效卷积神经网络EfficientNetV2, 并对EfficientNetV2-S的网络架构进行详细介绍, 3.2节对我们改进的算法模型进行介绍.

3.1 EfficientNetV2介绍

EfficientNetV2是由Tan等[19]在2021年发布的一组基于CNN的主干特征提取网络. EfficientNetV2网络不仅在准确率方面达到了当时的最高水平, 而且训练速度更快、参数数量更少. EffilcientNetV2系列共有S、M、L这3个不同的模型, 其中EfficientNetV2-S结构如表3所示.

EfficientNetV2-S中共有8个stage, Conv3×3是指3×3卷积块, Fused-MBConv1和Fused-MBConv4分别指Fused-MBConv中通道数为1和4, k3×3指卷积核的尺寸为3×3, SE0.25指SE (Squeeze-and-Excitation)模块中全连接层(Fully Connected, FC)中的节点数占特征矩阵的0.25. Pooling是池化层. 另外, EfficientNetV2-S是基线网络. M、L在S的模型框架基础上增加了网络复杂度, 在提升网络性能的同时, 仅增加了较少的计算成本. EfficientNetV2与EfficientNet[24]系列相比, 除了使用MBConv (Mobile Inverted Bottleneck Convolution)[19], 还在网络浅层使用Fused-MBConv[19], 解决了在大尺寸的训练图像上训练速度慢的问题. 同时, EfficientNetV2引入渐进式学习策略, 能够根据训练图像的尺寸动态调整正则化方法. 该策略不仅能够提升准确率, 同时还能提升训练速度. Fused-MBConv相较于MBConv减少了Depthwise卷积块, 其详细结构如图2所示. 其中, "H, W, 4C"指将输入卷积层的特征矩阵的通道数变为4倍, "H, W, C"指不做任何扩充.

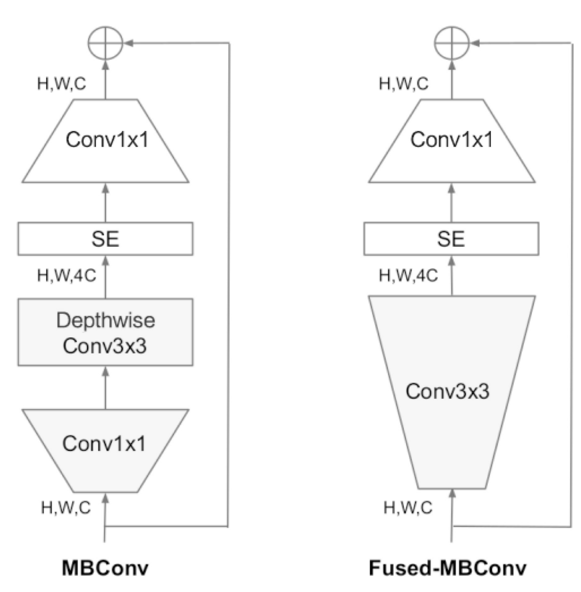


图 2 MBConv和Fused-MBConv的结构[19]

Fig. 2 Structures of MBConv and Fused-MBConv[19]

表 3 EfficientNetV2-S的结构

Table 3 EfficientNetV2-S structure

Stage	Operator	Stride	#Channels	#Layers
0	Conv3×3	2	24	1
1	Fused-MBConv1, k3×3	1	24	2
2	Fused-MBConv4, k3×3	2	48	4
3	Fused-MBConv4, k3×3	2	64	4
4	MBConv4, k3×3, SE0.25	2	128	6
5	MBConv6, k3×3, SE0.25	1	160	9
6	MBConv6, k3×3, SE0.25	2	272	15
7	Conv1 × 1 & Pooling & FC	-	1792	1

3.2 改进后模型架构

EfficientNets经历了从EfficientNet到EfficientNetV2的更新, EfficientNet包括B0至B7共8个模型, 它们具有相同的结构框架. EfficientNet-B0是基线网络, 其他均在输入图像分辨率、网络深度方面作了扩展, 以方便处理不同大小的图片. 我们在EfficientNet的8个模型和EfficientNetV2的3个模型上分别进行实验, 选出了在该星系图像数据集上表现最好的模型EfficientNetV2-S. 并将其与Alex-Net[25]、ResNet-34[26]、MobileNet-V2[27]和RegNet[28]等经典卷积神经网络进行对比, 前者表现出更好的分类性能. 因此我们考虑对模型EfficientNetV2-S进行进一步改进.

EfficientNetV2-S中使用的注意力机制模块是SE[29]模块, SE模块是对特征图进行二维全局池化操作, 通过压缩操作将高维的全局特征图转化为低维的特征向量, 以此提取通道级的全局特征, 进而用多层感知机进行非线性特征变换. 其更多关注通道间的信息, 而忽略了空间信息.

为了让模型能够结合不同注意力机制的优点, 进一步增强模型的特征提取能力, 我们考虑以下3种加入注意力模块的方法:

(1)在表3的stage0的3×3卷积后分别加入SE、CBAM (Convolutional Block Attention Module)[30]、ECA (Efficient Channel Attention)[31]、CA (Coordinate Attention)[32]和Triplet注意力模块.

(2)将stage1-6的SE换成CBAM、ECA、CA和Triplet模块.

(3)在stage7的1×1卷积后分别加入SE、CBAM、ECA、CA和Triplet注意力模块.

SE、ECA为通道注意力机制, CBAM、CA和Triplet是通道和空间信息相结合的注意力机制. 其中, Triplet注意力机制的创新点在于可以获取跨维度的交互信息来计算注意力权重, 同时建立起维度之间的依存关系, 而对通道信息和空间信息进行编码的计算开销是很小的, 可以忽略[20]. 以上都得益于Triplet重要的三分支结构. 这样, 卷积层的输出完成了通道与空间信息的跨通道交互, 在增强模型表征能力的同时, 还能抑制无效信息的干扰, 提升模型的鲁棒性. Triplet的结构如图3所示. 其中, 第1个分支用于实现通道信息C和高度信息H的跨维度交互, Z-Pool指最大池化或平均池化, Conv指7×7卷积层; 第2个分支用于实现通道信息C和宽度信息W的跨维度交互; 第3个分支最后对特征进行重排序. Avg指对3个分支输出特征进行相加求平均值.

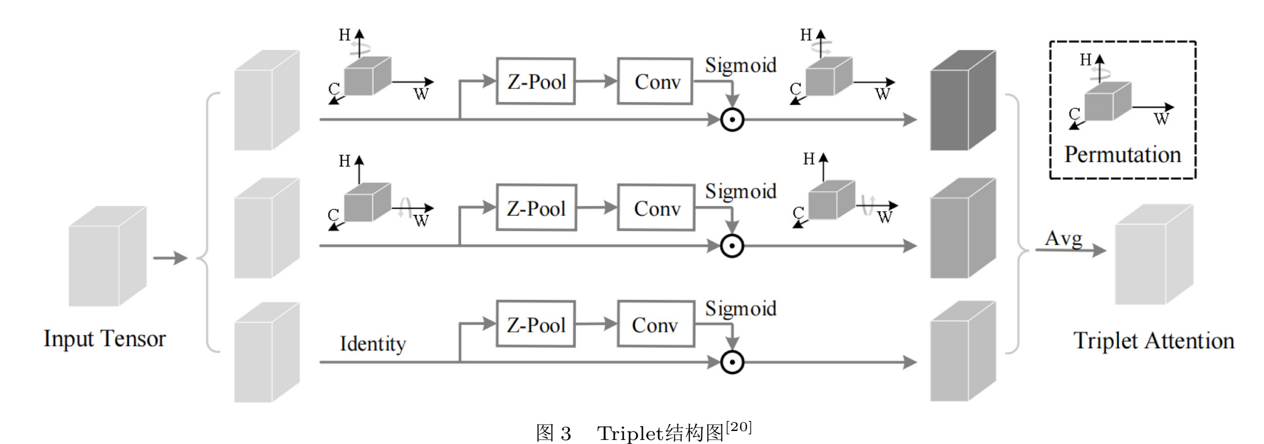


图 3　Triplet结构图[20]

Fig. 3　Structure of Triplet[20]

更进一步, 我们将在EfficientNetV2-S中stage7的1×1卷积层后加入Triplet注意力模块的模型称为EfficientNetV2-S-Triplet7. 图4展示了改进前EfficientNetV2-S中stage7的结构图, 图5展示了改进后EfficientNetV2-S-Triplet7中stage7 的结构图.

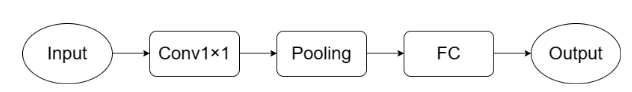


图 4　EfficientNetV2-S中stage7的结构图

Fig. 4　Structure of stage7 in EfficientNetV2-S

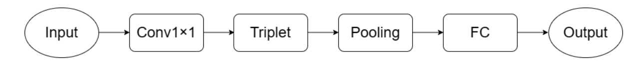


图 5　EfficientNetV2-S-Triplet7中stage7的结构图

Fig. 5　Structure of stage7 in EfficientNetV2-S-Triplet7

4　实验过程与结果

4.1　评价标准

在分类问题中, 有关样本真实类别与模型预测类别的组合的混淆矩阵如表4所示, 其他评估分类

模型的性能指标表示如下[33].

准确率:

$$\mathrm{Accuracy}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}\,; \tag{1}$$

查准率:

$$\mathrm{Precision}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\,; \tag{2}$$

查全率:

$$\mathrm{Recall}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}\,; \tag{3}$$

F1分数:

$$\mathrm{F1-score}=\frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}}\,. \tag{4}$$

ROC曲线(Receiver Operating Characteristic Curve)与AUC (Area Under Curve): 在预测任务中, f是分类学习器, 对于一个样本数量为m样例集$\boldsymbol{D}=\{(x_1,y_1),(x_2,y_2),\cdots,(x_m,y_m)\}$, 其中$y_i$是$x_i$示例的真实标记. ROC曲线是接收器操作特性曲线. AUC为ROC曲线下的面积, AUC可估算为:

$$\mathrm{AUC}=\frac{1}{2}\sum_{i=1}^{m-1}(x_{i+1}-x_i)\times(y_i+y_{i+1})\,. \tag{5}$$

表 4 分类结果混淆矩阵

Table 4 The confusion matrix of classification result

Predict label / True label	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

4.2 实验环境及参数

本文所用计算机的硬件环境为Intel (R) Xeon (R) CPU E5-2640 v4 @ 2.40 GHz处理器, 32 GB内存, NVIDIA GeForce RTX 3080 Ti显卡, 软件环境为64位Windows10操作系统. 所有实验均基于深度学习框架Pytorch, 所用开发语言为Python3.8.

在训练过程中, 我们采用随机初始权重. 首先采用交叉熵损失(Cross Entropy Loss)函数计算损失值; 然后采用SGD (Stochastic Gradient Descent)算法优化网络参数, 动量为0.9, 权重衰减为0.0001, 使用余弦退火衰减学习率, 批处理大小设置为32.

4.3 实验结果

针对2.2节提到的样本不平衡问题, 我们通过随机旋转、随机翻转、色彩畸变等方法对无核球侧向星系和扰动星系进行数据增强策略, 并基于网络模型AlexNet、 ResNet、 EfficientNet-B0和EfficientNetV2-S进行实验对比, 实验结果如表5所示. 从表5数据可以看出, 经过数据增强后, AlexNet、ResNet、EfficientNet-B0和EfficientNetV2-S的测试精度分别提高了1.53%、1%、0.82%以及1.11%, 这验证了数据增强策略的有效性.

紧接着, 本文对EfficientNet-B0至B7、EfficientNetV2-S、M、L及其他经典深度学习模型进行训练和测试, 实验结果如表6所示. 其中, RegNetY是在RegNetX中插入SE机制得到的, 400MF比200-MF具有更复杂的结构. 表6结果显示, EfficientNetV2-S的4项评价指标都是最高的, 其测试准确率为88.64%, 查准率为89.45%, 查全率为89.65%, F1分数为89.52%. 基于此我们发现EfficientNetV2-S在本文的椭圆星系、有核球侧向星系、无核球侧向星系、旋涡星系、棒旋星系和扰动星系6分类任务上有更好的特征提取能力. 因此, 我们选择EfficientNetV2-S模型进行改进.

表 5 数据增强前后的分类结果对比

Table 5 Comparison of classification results before and after data augmentation

Model	Accuracy without data augmentation	Accuracy with data augmentation
AlexNet	84.55%	86.08%
ResNet	86.88%	87.88%
EfficientNet-B0	86.93%	87.75%
EfficientNetV2-S	87.53%	88.64%

表 6 各深度学习模型性能比较

Table 6 Performance comparison of various deep learning models

Model	Accuracy	Precision	Recall	F1-score
EfficientNet-B0	87.75%	88.42%	88.52%	88.40%
EfficientNet-B1	87.93%	88.99%	88.43%	88.65%
EfficientNet-B2	88.07%	88.99%	88.83%	88.88%
EfficientNet-B3	88.33%	88.78%	89.50%	89.04%
EfficientNet-B4	88.00%	89.23%	88.62%	88.89%
EfficientNet-B5	87.97%	88.47%	88.98%	88.68%
EfficientNet-B6	88.56%	89.19%	89.60%	89.33%
EfficientNet-B7	88.01%	88.60%	89.11%	88.80%
EfficientNetV2-S	**88.64%**	**89.45%**	**89.65%**	**89.52%**
EfficientNetV2-M	88.32%	89.17%	89.61%	89.37%
EfficientNetV2-L	87.98%	88.51%	88.88%	88.66%
AlexNet	86.08%	86.33%	86.48%	86.61%
ResNet-34	87.88%	89.14%	88.73%	88.88%
MobileNetV2	86.85%	87.58%	87.64%	87.58%
RegNetX-200MF	83.69%	84.56%	84.09%	84.30%
RegNetY-400MF	85.41%	86.26%	85.55%	85.84%

4.4 EfficientNetV2-S改进模型的实验结果与分析

为了使EfficientNetV2-S的改进策略更加有效, 我们考虑在EfficientNetV2-S算法的stage0、stage7嵌入SE、CBAM、ECA、CA和Triplet模块, 或者将stage1-6中的SE模块换成其他4种注意力模块.

实验结果如表7所示. 表中模型EfficientNetV2-S-SE0、EfficientNetV2-S-CBAM0、EfficientNetV2-S-ECA0、EfficientNetV2-S-CA0和EfficientNetV2-S-Triplet0分别代表在stage0的3×3卷积后分别加入SE、CBAM、ECA、CA和Triplet模块; EfficientNetV2-S-CBAM、EfficientNetV2-S-ECA、EfficientNetV2-S-CA和EfficientNetV2-S-Triplet分别代表将EfficientNetV2-S中stage1-6的SE注意力机制换成CBAM、ECA、CA和Triplet模块; 模型EfficientNetV2-S-SE7、EfficientNetV2-S-CBAM7、EfficientNetV2-S-ECA7、EfficientNetV2-S-CA7和EfficientNetV2-S-Triplet7分别代表在stage7的1×1卷积后分别

加入SE、CBAM、ECA、CA和Triplet模块. 我们发现, 在stage0后加入5种注意力机制的的模型分类准确率都有不同程度的下降. 将stage1-6的SE换成CBAM、ECA注意力机制的模型分类准确率分别上升了0.04%和0.31%. 而对于在stage7的1×1卷积后加入5种注意力机制的模型来说, 只有加入CBAM模块的模型准确率下降了0.47%, 其他均有提升. 其中在stage7的1×1卷积后加入Triplet模块的模型达到89.03%的准确率、90.21%的查全率、89.93%的F1分数, 高于其他各种改进模型. 同时, EfficientNetV2-S-Triplet7与未改进的EfficientNetV2-S相比准确率提高了0.39%.

表 7 在EfficientNetV2-S插入注意力模块的实验结果
Table 7 Experimental results of inserting attention module in EfficientNetV2-S

Model	Accuracy	Precision	Recall	F1-score
EfficientNetV2-S-SE0	88.49%	89.42%	89.39%	89.37%
EfficientNetV2-S-CBAM0	88.27%	89.01%	89.22%	89.07%
EfficientNetV2-S-ECA0	88.29%	88.95%	89.37%	89.13%
EfficientNetV2-S-CA0	87.43%	88.55%	88.25%	88.38%
EfficientNetV2-S-Triplet0	88.10%	88.80%	89.09%	88.92%
EfficientNetV2-S-CBAM	88.68%	89.52%	89.97%	89.72%
EfficientNetV2-S-ECA	88.95%	**89.87%**	89.90%	89.87%
EfficientNetV2-S-CA	88.31%	88.93%	89.58%	89.22%
EfficientNetV2-S-Triplet	88.15%	88.92%	89.25%	89.05%
EfficientNetV2-S-SE7	88.87%	89.81%	89.88%	89.24%
EfficientNetV2-S-CBAM7	88.17%	89.08%	89.11%	89.08%
EfficientNetV2-S-ECA7	88.82%	89.56%	89.85%	89.69%
EfficientNetV2-S-CA7	88.69%	89.52%	89.68%	89.56%
EfficientNetV2-S-Triplet7	**89.03%**	89.69%	**90.21%**	**89.93%**

此外, 由于EfficientNetV2-S-Triplet7本身深度较大, 模型复杂度高, 与传统的ResNet-34等相比, 训练时间方面没有明显优势. 但是, 我们训练EfficientNetV2-S-Triplet7所有层的权重, 完成一个epoch的训练时间约为8.5 min, 快于EfficientNet系列中分类准确率最高的EfficientNet-B6. 因此在考虑分类性能的前提下, 我们所建立模型的训练时间是可以接受的.

为了清晰直观地展示我们构建模型的测试结果, 我们结合了数目混淆矩阵和百分比混淆矩阵. 图6展示了EfficientNetV2-S-Triplet7在测试集上的混淆矩阵. 图6 (左)是数目混淆矩阵, 图6 (右)是百分比混淆矩阵. 两个图的列均表示真实标签, 行均表示预测标签. 其中, 2098个椭圆星系、837个有核球侧向星系、939个无核球侧向星系、1971个旋涡星系、1389个棒旋星系和1113个扰动星系被正确分类. 204个旋涡星系被错分成椭圆星系, 占所有旋涡星系的8.50%, 我们在图7中展示了6个示例; 138个椭圆星系被分成旋涡星系, 占所有椭圆星系的6.03%, 我们在图8中展示了6个示例. 结合图7和图8我们观察发现, 二者易被错误分类是由一些旋涡星系的内部旋涡结构较模糊且边缘轮廓与椭圆星系相似所致.

测试集中6类星系模型的ROC曲线如图9所示. 每个颜色代表一个类, 真正例率(True Positive Rate, TPR)值越接近1, 假正例率(False Positive

Rate, FPR)值越接近0, 曲线预测效果越好, 即曲线越接近左上角, 曲线预测效果越好. 从图9中可以看出, 每一类的ROC曲线都表现得很好, 而旋涡星系的预测相对较差, 这是由于旋涡星系的图像分辨率较低、轮廓不明显, 另外, 旋涡星系与椭圆星系、棒旋星系、扰动星系的自身特征差异较小. 最后, EfficientNetV2-S-Triplet7模型的平均AUC值为0.9805, 表明该模型的整体预测性能良好.

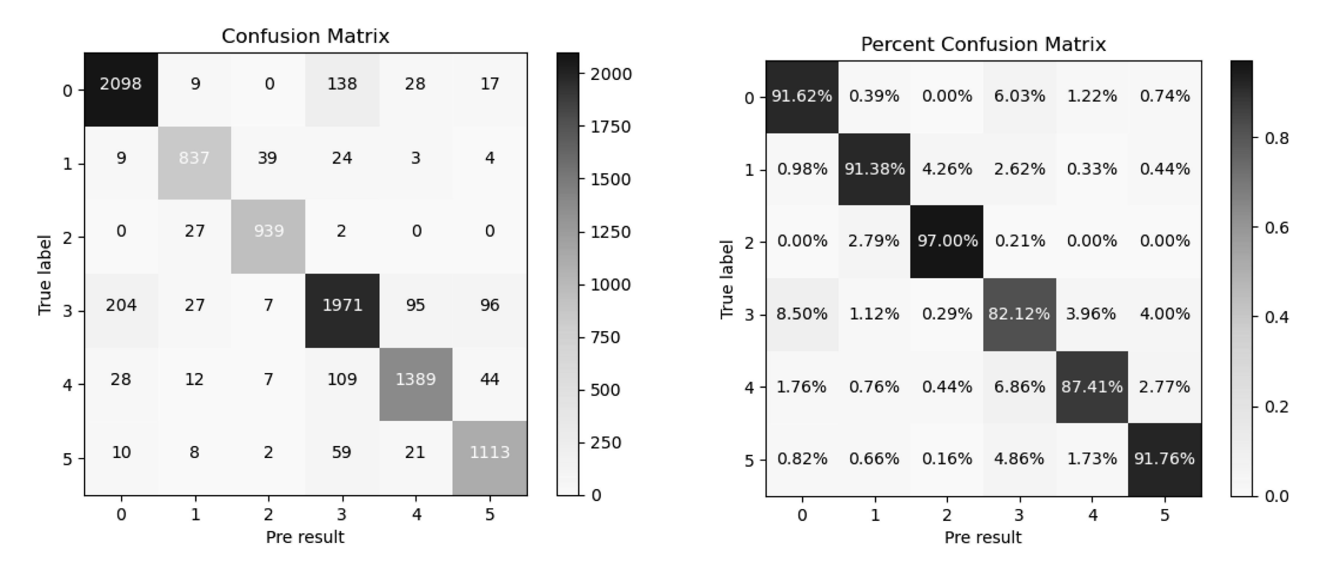


图 6 EfficientNetV2-S-Triplet7分类结果的数目混淆矩阵(左)和百分比混淆矩阵(右)

Fig. 6 Number confusion matrix (left) and percent confusion matrix (right) for EfficientNetV2-S-Triplet7 classification results

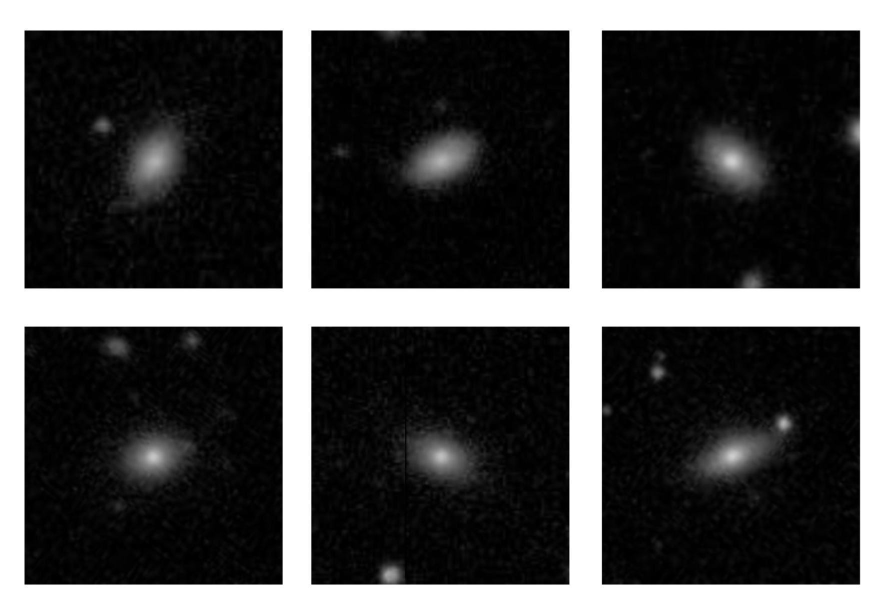


图 7 真实标签为旋涡星系, 预测结果为椭圆星系的图像示例.

Fig. 7 Example images where the true label is a spiral galaxy and the predict result is an elliptical galaxy.

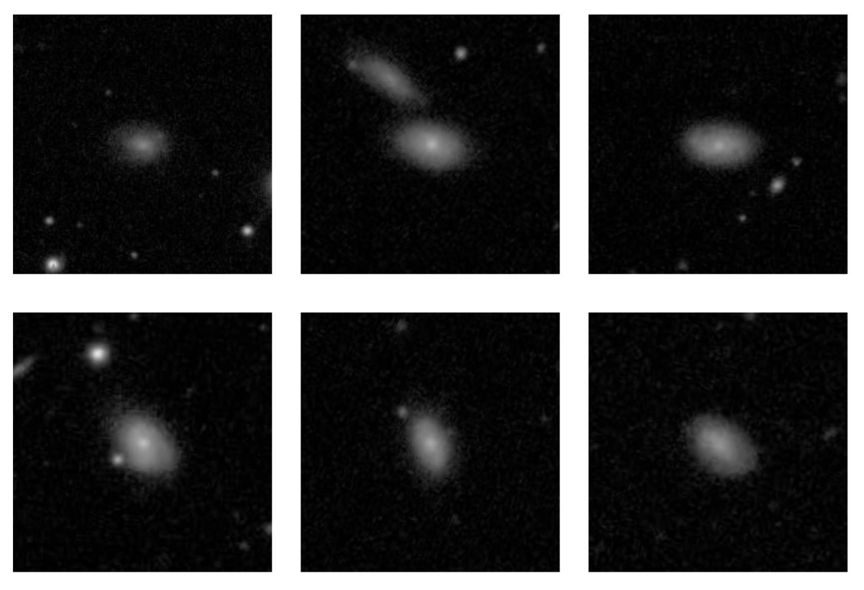


图 8 真实标签为椭圆星系, 预测结果为旋涡星系的图像示例.

Fig. 8 Example images where the true label is an elliptical galaxy and the predict result is a spiral galaxy.

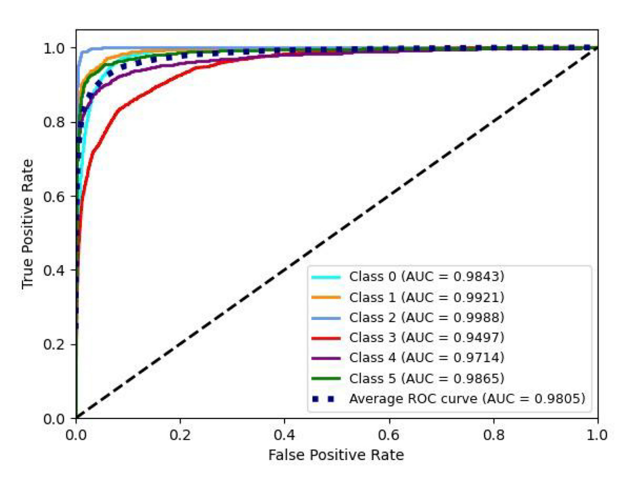


图 9 EfficientNetV2-S-Triplet7分类结果ROC曲线

Fig. 9 ROC for EfficientNetV2-S-Triplet7 classification results

5 结论

随着天文观测技术和设备的不断进步, 自动化分析海量天文数据越来越受到重视. 本文将EfficientNetV2网络应用于星系形态学分类数据集—第二期星系动物园中, 并与经典模型AlexNet、ResNet-34、MobileNetV2和RegNet进行对比, 实验表明, EfficientNetV2有更高的分类准确率. 其中EfficientNetV2-S在椭圆星系、有核球侧向星系、无核球侧向星系、旋涡星系、棒旋星系和扰动星系6分类任务上表现出最好的分类准确率、查准率、查全率和F1分数. 本文还采用了尺寸抖动、翻转、色彩畸变等图像增强手段来提高天文图像后期处理的质量, 然后在该模型的stage7的1×1卷积层后加入相对较优的Triplet注意力模块, 最终在93758张测试星系图片的准确率、查全率、F1分数方面均优于本文其他模型, 3项指标值分别为89.03%、90.21%、89.93%. 然而, 由于旋涡星系与椭圆星系、棒旋星系、扰动星系边缘相似, 错分比例较高. 同时, 星系测光图像分辨率与ImageNet传统图像相比分辨率更低, 分类比较困难. 因此, 下一阶段将致力于提高图像细粒度, 进一步增强网络的特征提取能力, 力争取得更好的分类效果. 另一方面, 本文数据来自地基望远镜, 其更易受到大气干扰的影响, 所以数据质量低于空间望远镜拍摄的数据. 那么将我们的模型应用于空间望远镜的数据中, 也可能取得较好的星系形态分类效果. 近年来, 一些学者致力于对DES (the Dark Energy Survey)[6]中的低表面亮度星系(Low Sur-

[6]https://www.darkenergysurvey.org/

face Brightness Galaxy, LSBG)进行形态分类, 在未来的工作中, 我们尝试基于本文模型解决该问题. 此外, GZ2与DeCaLS (the Dark Energy Camera Legacy Survey)[7]数据相似度较高, 因此我们考虑将本文改进的模型应用在DeCaLS的星系形态分类问题中, 并尝试更多类别的星系形态分类研究. 为了进一步保证模型的分类性能, 我们将在实现上述应用程序的同时优化其架构和超参数.

参 考 文 献

[1] Dieleman S, Willett K W, et al. MNRAS, 2015, 450: 1441
[2] 汪敏, 孔旭. 天文学进展, 2007, 25: 215
[3] Hubble E P. ApJ, 1926, 64: 321
[4] de Vaucouleurs G. ApJS, 1961, 5: 233
[5] Morgan W W. PASP, 1958, 70: 364
[6] Abraham R G, Valdes F, Yee H K C, et al. ApJ, 1994, 432: 75
[7] Abraham R G, van den Bergh S, Nair P. ApJ, 2003, 588: 218
[8] Naim A, Lahav O , Sodre L, et al. MNRAS, 1995, 275: 567
[9] Bazell D, Aha D W. AJ, 2001, 548: 219
[10] Owens E, Grifiths R, Ratnatunga K. MNRAS, 1996, 281: 153
[11] Calleja J, Fuentes O. MNRAS, 2004, 349: 87
[12] Cheng T Y, Conselice C J, Aragón-Salamanca A, et al. MNRAS, 2020, 493: 4209
[13] Zhang Z, Zou Z, Li N, et al. A&A, 2022, 22: 055002
[14] Kim E J, Brunner R J. MNRAS, 2016, 464: 4475
[15] 戴佳明, 佟继周. 天文学进展, 2018, 36: 384
[16] Goyal L M, Arora M, Pandey T, et al. EScIn, 2020, 13: 1427
[17] 王林倩, 邱波, 罗阿理, 等. 天文研究与技术, 2022, 19: 359
[18] 艾霖嫔, 徐权峰, 杜利婷, 等. 天文学报, 2022, 63: 42
[19] Tan M X, Le Q V. Proceedings of the 38th International Conference on Machine Learning. Dubai: PMLR, 2021, 139: 10096
[20] Misra D, Nalamada T, Arasanipalai A U, et al. Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2021: 3138
[21] Willett K W, Lintott C J, Bamford S P, et al. MNRAS, 2013, 435: 2835
[22] LeCun Y, Bottou L, Bengio Y, et al. Proceedings of the IEEE, 1998, 86: 2278
[23] Gu J, Wang Z, Kuen J, et al. PatRe, 2018, 77: 354
[24] Tan M X, Le Q V. Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019, 97: 6105
[25] Krizhevsky A, Sutskever I, Hinton G E. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc, 2012, 1: 1097
[26] He K M, Zhang X Y, Ren S Q, et al. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770
[27] Sandler M, Howard A, Zhu M, et al. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510
[28] Radosavovic I, Kosaraju R P, Girshick R, et al. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10425
[29] Hu J, Shen L, Sun G. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132
[30] Woo S, Park J, Lee J Y, et al. Proceedings of the 15th European Conference on Computer Vision. Munich: Spinger, 2018: 3
[31] Wang Q, Wu B, Zhu P, et al. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531
[32] Hou Q, Zhou D, Feng J. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13708
[33] 周志华. 机器学习. 北京: 清华大学出版社, 2016: 28-35

[7] https://www.legacysurvey.org/decamls/

EfficientNetV2-S-Triplet7: An Improved Algorithm for Morphological Galaxy Classification

ZHONG Zheng-di TU Liang-ping FENG Xue-qi LI Juan LI Xin
(School of Science, University of Science and Technology Liaoning, Anshan 114051)

Abstract The morphology of galaxies is closely related to the formation and evolution of galaxies, and its morphological classification is a significant part of the follow-up research of galaxy astronomy. With the emergence of massive astronomical observation data, the automatic analysis of astronomical data has attracted more and more attention. To solve this problem, the advanced deep learning backbone network EfficientNetV2 is utilized to analyze the effects of different attention mechanism types and usage nodes on network performance, and an improved algorithm model named EfficientNetV2-S-Triplet7 is constructed to realize automatic classification of galaxy morphology. More than 240 thousand photometric images from Galaxy Zoo 2 are used as initial data for experimental tests. In the process of data preprocessing, image enhancement methods such as size jittering, flipping and color distortion are adopted to solve the problem of image number imbalance. After conducting comparative experiments on the same series of classic and cutting-edge deep learning algorithms AlexNet, RegNet, MobileNetV2 and ResNet-34, it is concluded that the EfficientNetV2-S-Triplet7 algorithm has the best test results in classification accuracy, recall and F1-score. In 9375 test images, the three index values can reach 89.03%, 90.21% and 89.93%, respectively, and the precision can reach 89.69%, ranking the third among other models. The results show that EfficientNetV2-S-Triplet7 algorithm can be effectively applied to the morphological classification of large-scale galaxy data.

Key words techniques: image processing, methods: data analysis, methods: classification, galaxies: structure